Effects of verb-action congruence on sensorimotor processing of goal-directed actions in two-year-olds: A technical report

Katharina Antognini

Department of Psychology, University of Zurich, Switzerland

Moritz M. Daum

Department of Psychology, University of Zurich, Switzerland

Neuroscience Center Zurich, University of Zurich and ETH Zurich, Switzerland

Correspondence concerning this manuscript should be addressed to:

Katharina Antognini

University of Zurich

Department of Psychology

Developmental Psychology: Infancy and Childhood

Binzmuehlestrasse 14, Box 21

8050 Zurich

Switzerland

Email: k.antognini@psychologie.uzh.ch

Abstract

Language and action share a common processing system, namely the sensorimotor system. Sensorimotor activity is associated with action prediction and action-verb processing already early during verb acquisition. Action verbs can have a positive effect on action prediction, if the action verb matches the subsequently perceived action. However, it is yet unclear if semantic congruence between the action verb and the action drives this effect, or rather effector-limb congruence (i.e., both the action verb and the action imply an action that is, for instance, performed with the hand). The current study investigated whether semantic congruence between an action verb and an action, compared to semantic incongruence, has different effects on action perception. We presented two-year-olds with sentences comprising action verbs, which either corresponded semantically to a subsequently observed action or not. To assess sensorimotor activity we measured the suppression of the mu and the beta rhythm by means of electroencephalography (EEG). Results are mixed. On the one hand semantic congruence did not affect mu suppression during action perception in toddlers who had all action verbs in their expressive vocabulary. On the other hand, the group of toddlers who did not have all action verbs in their expressive vocabulary did show a difference in mu suppression. In contrast to the mu band, the beta band revealed a power difference during action perception for toddlers who had all action verbs in their expressive vocabulary, but not for the other group.

Keywords: mu rhythm, EEG, action prediction, predictive coding, mirror neuron system

Introduction

Actions and language are interrelated because they are both processed in a prediction system, which the sensorimotor system is part of (Antognini, Hauser, & Daum, submitted; Lupyan & Clark, 2015). Empirical studies have shown that the sensorimotor system is involved in processing observed actions as well as action verbs already early in development (Antognini & Daum, 2017; Marshall & Meltzoff, 2014). Action verbs are words that are used to describe perceivable/visible processes, states, and relations, while abstract verbs describe invisible states such as existing and believing (Golinkoff & Hirsh-Pasek, 2008). The sensorimotor system processes observed actions from the age of 8 months (Nyström, Ljunghammar, Rosander, & von Hofsten, 2011; Southgate, Johnson, Osborne, & Csibra, 2009), while action verbs undergo this processing around 18-24 months of age (Antognini & Daum, 2017; Antognini et al., submitted). In contrast, abstract verbs and pseudoverbs (i.e., invented verbs that are phonotactically correct) involve the sensorimotor system to a lesser extent or not at all (Antognini & Daum, 2017; Fargier et al., 2012; Moreno, de Vega, & León, 2013).

Studies in adults have demonstrated that the context, in which verbs are presented influences the way in which the verbs are processed, that is, sensorimotor processing of action-related language is highly context dependent (Lam, Bastiaansen, Dijkstra, & Rueschemeyer, 2016; Rueschemeyer, Brass, & Friederici, 2007; Rueschemeyer, Lindemann, van Rooij, van Dam, & Bekkering, 2015; Tomasino, Weiss, & Fink, 2010). Similarly, it has been shown that observed actions are processed differently depending on the language context they are presented in (Gampe & Daum, 2014; Springer, Huttenlocher, & Prinz, 2012; Springer & Prinz, 2010). Gampe and Daum (2014) have shown that two-year-olds anticipate the goal of a goal-directed means-end action more quickly if these actions are presented after a corresponding action verb, compared to a linguistically neutral expression. This suggests

that toddlers benefit from the information provided by the action verb to process the subsequently observed action. Recently, Antognini, Hauser, and Daum (submitted) have shown that the sensorimotor system is involved in this facilitation. Toddlers of 24 months of age activated their sensorimotor system more during action observation if the action was preceded by an acoustically presented action verb that corresponded to the action (e.g., "I'll show you drawing"), compared to a neutral expression (e.g., "I'll show you something"). This was interpreted a facilitation effect due to referential predictions (i.e., predictions based on the action verb about what will happen next; Fischer & Zwaan, 2008) onto subsequent action prediction (Antognini et al., submitted). However, the question remained how deeply the meaning of the action verb is processed by the sensorimotor system in order to elicit facilitation in action processing.

On the one hand, one can argue that the sensorimotor system is part of a prediction network, which processes meaning, thus semantics (Kilner, 2011). This network generates predictions from the inferred meaning (i.e., manner, path, and goal) and passes it along the network, including the sensorimotor system. On each level, the accuracy of the prediction is evaluated and errors are sent back to the higher level in order to check if the inferred meaning is plausible given the input (Kilner, 2011; Kilner, Friston, & Frith, 2007a).

On the other hand, literature on action verb processing suggests that action verbs are processed with respect to their effector limb, that is, hand verbs are processed in hand areas, while foot verbs are processed in foot areas (Hauk, Johnsrude, & Pulvermüller, 2004; James & Maouene, 2009; Tettamanti et al., 2005). Therefore, it is possible that a facilitation effect from the action verb onto the observed action takes place if the effector limbs correspond (e.g., hearing a hand verb, seeing a hand action) irrespective of semantic congruence.

The aim of the current study was to clarify the question whether congruence in semantics or congruence in effector limb elicits facilitated action processing. To this end, we presented 24-month-olds with sentences comprising action verbs. The overall design

corresponded to the procedure used in Antognini, Hauser, and Daum (submitted). After each sentence, the toddlers observed a video clip of either displaying the corresponding action or a different action, but effectuated with the same limb. To measure facilitated action processing, we applied electroencephalography (EEG) to record the sensorimotor activity via the suppression of the mu and the beta rhythm.

Both rhythms originate from the sensorimotor system and are harmonic to each other (Hari et al., 1998; Hari & Salmelin, 1997). Mu suppression is a widely used indicator for sensorimotor activity in adults, children and infants (Fox et al., 2016; Marshall, Bar-Haim, & Fox, 2002) and originates from somatosensory cortices (Hari & Salmelin, 1997). However, the relation between mu suppression and a particular function within the predictive coding framework (e.g., making predictions or processing prediction errors) is not well described.

Beta suppression has also been used with infants and toddlers in the past to study action processing, however less frequently than mu suppression (Liao, Acar, Makeig, & Deak, 2015; van Elk, van Schie, Hunnius, Vesper, & Bekkering, 2008). It has been suggested that the beta rhythm originates from the motor cortex (Hari & Salmelin, 1997) and that beta suppression reflects predictive motor preparation (Kilner, Bott, & Posada, 2005). However, more recently, modulations in the beta rhythm have been associated with model updating and processing prediction errors (Palmer, Zapparoli, & Kilner, 2016).

In sum, the mu rhythm and the beta rhythm originate from different sites within the sensorimotor system and seem to reflect separate processes within a predictive coding framework. Therefore, it is important to study modulations in both frequency bands when investigating action prediction as in the current study.

Materials and Methods

Participants

The final sample included 37 toddlers (19 female) between 23 and 24 months of age (M = 748.7 days, range = 722-761 days). An additional 5 toddlers (2 female) were tested but excluded because of refusal during the application of the EEG sensor net (n = 3) or failure to reach a minimum of 2 artifact-free trials per experimental condition (n = 2). It is common practice to use a lean trial criterion in infant studies (de Klerk, Johnson, Heyes, & Southgate, 2015; Southgate et al., 2009). All toddlers were born full term (gestation ≥ 37 weeks, birth weight ≥ 2500 g), grew up in monolingual Swiss German households, and were recruited from a database of parents who volunteered to participate in infant studies. Caregivers gave informed written consent, and the study was approved by the local ethics committee in accordance with the 1964 Helsinki declaration and its later amendments. For their participation, the toddlers received an age-appropriate toy (value equivalent to 5 USD) and a certificate of participation with a photograph of themselves.

Stimulus material

We used two types of stimuli: auditory and visual. Eight sentences, prerecorded by two Swiss-German native speakers (male and female), were used as auditory stimuli. Two speakers were prerecorded to enhance the variability of the stimuli because enhance variability increases word processing in toddlers (Rost & McMurray, 2009). Each speaker recorded four of the sentences, respectively. Each sentence started with „Ich zeig dir“ [I'll show you] (0-900 ms) and was completed by one of four different action verbs (950-1650 ms). The following early-acquired action verbs (Szagun, Stumper, & Schramm, 2009) were used in the verb condition: „maale“ [drawing], „schniide“ [cutting], „baue“ [building], and „zuemache“ [closing].

Four video clips served as visual stimuli. Each of the four video clips displayed a different means-end action corresponding to the action verbs from the auditory stimuli: drawing a spiral on a piece of paper, cutting a toy carrot into three pieces, building a tower with three building blocks, and closing the lid of a jar. Video clips displayed a female actress sitting at a table such that the torso, arms, and hands were visible. Two objects (means object, goal object) were placed on the table with counter-balanced positions (left, right). The action sequences were as follows (see *Figure 1*): Each video clip started with the actress sitting at the table with her hands resting on the table (0-400 ms). Then, she lifted the hand nearest to the means object to reach for it and grasped it ipsilaterally (400-1640 ms). She transported the means object to the goal object (1640-2880 ms), and performed the appropriate action on the goal object (2880-4120 ms). After action completion, hands were again resting on the table. In total, video duration was either 6000 ms for actions without a tool object (building and closing), or 7000 ms for actions with a tool object (drawing and cutting). In videos with a tool object, there was an additional action step (placing tool on table) before the resting hands in the end.

Procedure

Prior to the EEG measurement, the caregivers completed a language questionnaire, which was an adapted version of the German McArthur CDI (Szagun et al., 2009), to indicate which verbs, out of one hundred, their child already actively produced in daily life (i.e., spontaneous utterance and consistent use to describe an action).

EEG measurements took place in an electrically shielded, sound-attenuated, and dimly lit room. The toddler was seated on the caregiver's lap at approximately 60 cm distance from a 17-inch computer screen with adjacent loudspeakers. The stimuli were presented with Psychtoolbox (version 3.0.11) in MATLAB R2013a. During the first 1000 ms of each trial, a white fixation dot was presented on the screen. After that, a sentence with one of four action

verbs was presented via the loudspeakers while no visual stimulation was present. Following the sentence, a video clip with either a corresponding means-end action (congruent condition) or a non-corresponding means-end action (incongruent condition) was displayed on the screen. We presented a maximum of 8 blocks with 16 trials each (4 female speaker/congruent, 4 male speaker/congruent, 4 female speaker/incongruent, 4 male speaker/incongruent). Between blocks, a one-minute play break was introduced to keep the toddler's interest for the experimental trials as long as possible. The experiment was terminated when the toddler was no longer attentive (i.e., looking away from the screen, crying, getting out of the caregiver's lap), or when the maximum of 8 blocks was reached. Video cameras recording the lateral and front view were used to monitor the toddlers throughout the experiment. The experimenter released an attention grabbing video (bouncing cube with sound effect) each time the toddler moved or looked away from the screen.

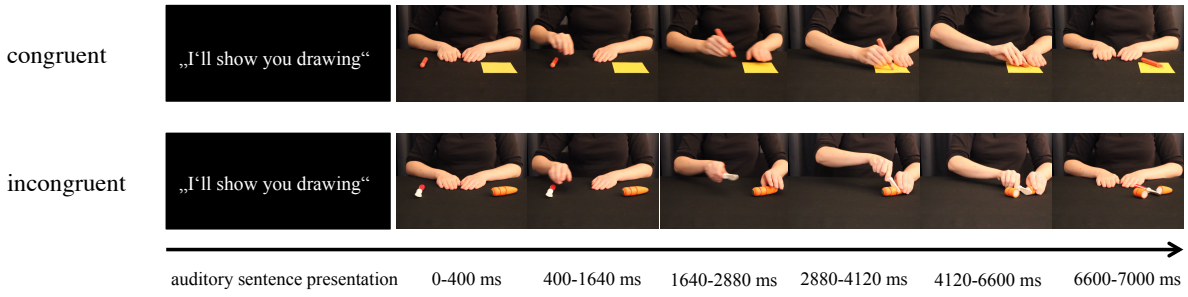

Figure 1. Exemplary composition of two experimental trials. A video clip, either corresponding to the meaning of the action verb (person drawing on a piece of paper; congruent condition) or not (person cutting a carrot; incongruent condition), followed the sentence with the action verb. The structure was analogous for all actions used in the paradigm.

EEG recording and analysis

The EEG measured with a 128-electrode HydroCel sensor net (EGI, Eugene, OR, USA) and a NetAmps 300 amplifier (EGI, Eugene, OR, USA) at a sampling rate of 500 Hz. During recording, data was referenced to the vertex electrode (Cz) and high-pass filtered at 0.1 Hz. Impedances were kept below 50 kΩ. We performed offline preprocessing in the EEGLAB toolbox (Delorme & Makeig, 2004): band-pass filtering between 0.3-30 Hz; removal of outermost electrodes per default; removal of bad electrodes after visual inspection; independent component analysis (ICA) to remove artifacts due to eye blinks, saccades, sweating, and heart beat (Delorme & Makeig, 2004); spherical interpolation of missing channels; and rereferencing data to common average reference.

Continuous EEG data was segmented into epochs from -1000 ms to 6400 ms with respect to sentence onset. These epochs were chosen to analyze only EEG trials for which the toddler attended the sentence and the video attentively and quietly. Trials in which the toddlers moved or did not look at the screen (during either audio or video presentation) were removed from further analysis. The toddlers contributed on average 12.7 artifact-free trials to the congruent condition (range 3-24) and 11.6 artifact-free trials to the incongruent condition (range 4-22).

We applied a wavelet analysis with 7 constant cycles between 4 and 20 Hz to the segmented data, using Matlab (R2014b). The mu band in 24-month-olds peaks at 8 Hz (Berchicci et al., 2011). Therefore, we selected a frequency band of interest between 6 and 10 Hz. The frequency band for the beta rhythm was chosen at 15-18 Hz according to previous literature (Liao et al., 2015). Event-related desynchronization (ERD) values were extracted according to Pfurtscheller (2001). We selected the period during sentence presentation in which the action verb was presented as baseline (950-1650 ms after sentence onset). This baseline was chosen because it is equal in both conditions and directly precedes the time

window of interest. A further reason was that this baseline could account for any possible but unexpected differences during action-verb presentation that could potentially propagate to the action observation period. By choosing the action verb as baseline, we correct for this potential propagation. ERD was calculated for the time period during which the grasp towards the tool object in the video clip was effectuated (0-1640 ms after video onset; grasping time-window). The grasping time-window was based on the time window chosen by Gampe & Daum (2014). We selected three different electrode clusters of interest: left centro-parietal (E30, E31, E36, E37, E42, E53, E54) and right centro-parietal (E79, E80, E86, E87, E93, E104, E105), and occipital (E66, E69, E70, E71, E74, E75, E76, E82, E83, E84, E89). The occipital electrode cluster served to account for potential attentional effect. Statistics were computed in the R statistical package (R Core Team, 2017).

Results

It has previously been shown that toddlers differ in the sensorimotor processing of observed actions depending on their expressive verb vocabulary (Gampe & Daum, 2014). Additionally, it has been suggested that sensorimotor activity is experience-dependent (Kilner et al., 2007a; Kilner, Friston, & Frith, 2007b). Therefore, we chose to analyze two separate groups of toddlers. The first group consisted of toddlers with all the action verbs from the paradigm in their expressive vocabulary (proficient group, $N = 18$), according to parental report in the questionnaire. The second group consisted of toddlers who did not have all action verbs in their expressive vocabulary (less proficient group, $N = 19$).

Difference in mu power during action observation

To account for the dynamics in power change of the mu rhythm (as suggested by Fox et al., 2016), we applied one paired *t*-test per time point ($n = 821$) on the ERD values [%] of the mu frequency band (6-10 Hz) within the grasping time-window (0-1640 ms ms after video

onset) to compare the ERD during action observation in both conditions. Because the tested power values within the time window were continuous, and therefore not independent from each other, we applied Fisher's Omnibus test with permutation statistics (iterations $N = 10'000$) as suggested by Potter and Griffiths (2006). With this procedure, the *p*-values from the single *t*-tests are combined into an overall *p*-value with Fisher's function (Fisher, 1932).

For the proficient group, results indicated no significant differences in ERD during the observation of grasping between the congruent and incongruent condition: left centro-parietal cluster, $\chi^2(N = 18) = 568.93, p = 1$; right centro-parietal cluster, $\chi^2(N = 18) = 651.33, p = 1$; occipital cluster, $\chi^2(N = 18) = 1292.54, p = 1$. For the less proficient group, the analysis indicated a significant difference between conditions in the left centro-parietal, $\chi^2(N = 19) = 1864.04, p < .001$, the right centro-parietal, $\chi^2(N = 19) = 2040.86, p < .001$, and in the occipital cluster, $\chi^2(N = 19) = 2340.85, p < .001$. In the left centro parietal cluster, ERD was greater in the congruent compared to the incongruent condition (see *Table 1*). In contrast, in the right-centro parietal cluster, ERD was greater for the incongruent than the congruent condition (see *Table 1*). In analogy, in the occipital cluster, ERD was greater for the incongruent compared to the congruent condition (see *Table 1*).

Difference in beta power during action observation

This analysis investigated the differences between conditions in terms of ERD values of beta frequency band (15-18 Hz, Liao et al., 2015) within the grasping time-window (0-1640 ms after video onset) to compare the ERD during action observation in both conditions. Except for the frequency band of interest, the analysis was analogous to the previous section.

For the proficient group, results indicated significant differences in ERD between the conditions, in the left centro-parietal, $\chi^2(N = 18) = 2300.75, p < .001$, and in the occipital cluster, $\chi^2(N = 18) = 3104.02, p < .001$. In both clusters ERD was enhanced for the congruent

compared to the incongruent condition (see *Table 1*). There were no significant differences in ERD in the right centro-parietal cluster, $\chi^2(N = 18) = 1220.01, p = 1$. For the less proficient group, results indicated no significant differences in any electrode cluster: left centro-parietal, $\chi^2(N = 19) = 1428.07, p = .999$; right centro-parietal, $\chi^2(N = 19) = 1340.31, p = .999$; occipital, $\chi^2(N = 19) = 1339.98, p = .999$.

Table 1

Mean ERD/ERS during Action Observation for both Groups of Toddlers

	Mu ERD		Beta ERD	
Cluster	Congruent	Incongruent	Congruent	Incongruent
		Proficient		
Left centro-parietal	-0.46 (2.91)	-0.89 (3.43)	**-0.27 (3.95)**	**2.87 (4.40)**
Right centro-parietal	-2.72 (2.83)	-3.71 (3.63)	0.60 (3.97)	1.35 (3.82)
Occipital	-3.21 (3.10)	-1.62 (3.25)	**-1.52 (3.44)**	**2.38 (3.10)**
		Less proficient		
Left centro-parietal	**-1.35 (4.52)**	**0.33 (5.37)**	-1.86 (3.65)	-2.06 (3.07)
Right centro-parietal	**-3.44 (4.47)**	**-5.10 (5.26)**	-2.54 (3.01)	-2.14 (2.54)
Occipital	**0.33 (5.25)**	**-2.08 (4.83)**	-2.53 (3.95)	-2.35 (3.28)

Note. The table is divided by the two groups of toddlers (proficient, less proficient). Given are the mean ERD (negative) or ERS (positive) values. Standard deviations (*SD*) are given in the brackets. Significant differences are printed in bold face.

Discussion

This study investigated the role of semantic action-verb congruence in subsequent action processing. To this end, we recorded mu and beta suppression of two-year-olds by means of EEG during action verb presentation and action observation. In order to account for experience-dependency, the toddlers were split in two groups (proficient, less proficient), which were analyzed separately. Results were mixed and will be discussed in the following sections. Please note that this discussion is not finalized. We would appreciate the readers to share their thoughts about this discussion. The current manuscript is a technical report, which is intended to inform future studies on the topic.

Mu and beta suppression during action processing

An overview of the results can be found in Table 2. The reader might appreciate the differences between the two groups of toddlers (proficient, less proficient) in terms of the modulation of the two frequency bands (mu and beta). Roughly, the proficient group only showed modulations of the beta band during action observation, whereas the less proficient group displayed modulations only in the mu band. These results will be discussed in more detail below.

Table 2

Direction of ERD Differences during Action Observation for both Groups of Toddlers

Cluster	Mu	Beta
	Proficient	
Left centro-parietal	*n.s.*	congruent < incongruent ***
Right centro-parietal	*n.s.*	*n.s.*
Occipital	*n.s.*	congruent < incongruent ***
	Less proficient	
Left centro-parietal	congruent < incongruent ***	*n.s.*
Right centro-parietal	incongruent < congruent ***	*n.s.*
Occipital	incongruent < congruent ***	*n.s.*

Note. n.s. = non significant

*** $p < .001$

For the mu band, we found differences in the ERD between conditions for the less proficient group only. In the left centro-parietal cluster the congruent condition resulted in enhanced mu ERD compared to the incongruent condition. A previous study showed similar findings, indicating enhanced left-central mu ERD in the congruent condition compared to a neutral condition, in which a neutral linguistic stimulus was presented before the video clip (Antognini et al., submitted). However, in this previous study, the difference between conditions could only be shown for the proficient group but not for the less proficient group (Antognini et al., submitted). These divergent findings could be due to the different conditions used in the two experiments (neutral/congruent in Antognini et al., submitted vs. congruent/incongruent in the present study). Nevertheless, the current findings might demonstrate that the less proficient group used the semantics of the congruent action verb to

make referential predictions about the video clip that will follow (Antognini et al., submitted; Fischer & Zwaan, 2008). However, only if this expectation was matched, it elicited enhanced action prediction through a referential prediction for the congruent video clip. A different interpretation could be that action prediction was rather suppressed in the incongruent video clip compared to the congruent video clip, due to the incongruence of the action verb and the video clip. It is not possible to decide which of the two interpretations is more plausible because we do not have measures of action prediction for the video clips in a case in which they were not preceded by an action verb. However, ERD values indicate that the incongruent condition was characterized by an event-related synchronization (ERS) near zero, while the congruent condition actually showed ERD (see Table 1). ERS has been identified as an indicator for inhibition (Pfurtscheller, 1992, 2001), which sustains the assumption that action prediction could have been suppressed by through the incongruence between action verb and observed action.

The less proficient group also demonstrated condition differences with respect to the mu band in the right centro-parietal and in the occipital cluster. Interestingly, results went into a different direction than for the left centro-parietal cluster (see Table 2). In the right centro-parietal cluster and in the occipital cluster, mu ERD was enhanced for video clip, which was incongruent with the action verb, compared to the video clip, which was congruent. This finding potentially points to an increased allocation of attention in the incongruent condition (Pfurtscheller, 2001; Yoo, Cannon, Thorpe, & Fox, 2016). However, this does not fully explain the condition difference in the right centro-parietal cluster, unless we interpret the right centro-parietal differences as pure topographical spreading from occipital sites.

Although these explanations might seem plausible, it is still unclear why the group with less experience with action verbs shows differences in mu ERD during action observation, while the more experienced-group doesn't. We would have expected the groups to differ in the opposite direction because expertise with a stimulus has been shown to enhance

sensorimotor responses in adults and infants (Cannon et al., 2014; Liew, Sheng, Margetis, & Aziz-Zadeh, 2013). However, it might be that mu suppression is not the best indicator in this paradigm because we are interested in possible error-processing mechanisms and updating of predictions. It has been suggested that the beta rhythm is a good indicator for investigating prediction errors and model updating in a predictive framework (Palmer et al., 2016). The assumption is that the beta rhythm displays suppression under conditions of sensory uncertainty, thus high updating demands. Furthermore, the beta band is supposed to synchronize when uncertainty is small, thus to inhibit further calculation of prediction errors (Palmer et al., 2016). Therefore, we will now discuss the results from the beta band.

In the beta band, we observed condition differences only in the proficient group, that is, in toddlers with expressive experience with the action verbs that we presented. Results indicated beta ERD for the congruent condition in the left centro-parietal and occipital cluster, while ERS was present for the incongruent conditions in the same clusters (see Table 2). In the congruent condition, the video clip matches the referential prediction based on the action verb. It might be the case that under these circumstances, constant model updating (comparison between referential prediction and action prediction) is necessary until the video clip has finished. According to Palmer and colleagues (2016) constant model updating under persisting uncertainty results in beta ERD, which is in line with our findings. In contrast, in the incongruent condition, model updating might be terminated earlier because the incongruence has been detected. Therefore, the uncertainty does not persist, which is associated with beta ERS, according to Palmer and colleagues (2016). Again, this is in line with our findings. Therefore, one might argue that sensorimotor activity during action perception is associated with processing the semantic congruence of a previously presented action verb. However, the processing of semantic congruence is rather reflected in the beta band than in the mu band. Importantly, it might be the case that the less proficient group is more in need of constant model updating in both conditions because they cannot equally

benefit from the information given by the action verb compared to the proficient group. This hypothesis is in line with our findings, which show beta ERD in both conditions for the less proficient group, instead of beta ERS for the incongruent condition (see Table 1).

Conclusion

To conclude, we are aware of the fact that the current study leaves many open questions and potentially unresolvable discrepancies. Therefore, we would appreciate if the readers of this manuscript contact us with their thoughts and ideas. The aim of this manuscript was to provide the information concerning this experiment to improve future work in the field of early neural interrelations between action and language.

Acknowledgements

During the work on her dissertation, Katharina Antognini was a pre-doctoral fellow of LIFE (International Max Planck Research School on the Life Course; participating institutions: MPI for Human Development, Humboldt-Universität zu Berlin, Freie Universität Berlin, University of Michigan, University of Virginia, University of Zurich). Furthermore, we thank Marco Bleiker for recruitment and organization, and Sarah Hauser for her help with the EEG measurements.

References

Antognini, K., & Daum, M. M. (2017). Toddlers show sensorimotor activity during auditory verb processing. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2017.07.022

Antognini, K., Hauser, S., & Daum, M. M. (submitted). Action verbs facilitate sensorimotor prediction of observed actions in toddlers. *Developmental Cognitive Neuroscience*.

Berchicci, M., Zhang, T., Romero, L., Peters, A., Annett, R., Teuscher, U., … Comani, S. (2011). Development of mu rhythm in infants and preschool children. *Developmental Neuroscience*, *33*(2), 130–143. https://doi.org/10.1159/000329095

Cannon, E. N., Yoo, K. H., Vanderwert, R. E., Ferrari, P. F., Woodward, A. L., & Fox, N. A. (2014). Action experience, more than observation, influences mu rhythm desynchronization. *PloS One*, *9*(3), e92002. https://doi.org/10.1371/journal.pone.0092002

de Klerk, C. C. J. M., Johnson, M. H., Heyes, C. M., & Southgate, V. (2015). Baby steps: Investigating the development of perceptual–motor couplings in infancy. *Developmental Science*, *18*(2), 270–280. https://doi.org/10.1111/desc.12226

Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., & Nazir, T. A. (2012). Learning to associate novel words with motor actions: Language-induced motor activity following short training. *Cortex*, *48*(7), 888–899. https://doi.org/10.1016/j.cortex.2011.07.003

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, *61*(6), 825–850. https://doi.org/10.1080/17470210701623605

Fisher, R. A. (1932). *Statistical methods for research workers* (4th ed., enlarged). Edinburgh: Oliver & Boyd.

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., … van IJzendoorn, M. H. (2016). Assessing human mirror activity with EEG mu rhythm: A meta-analysis. *Psychological Bulletin*, *142*(3), 291–313. https://doi.org/10.1037/bul0000031

Gampe, A., & Daum, M. M. (2014). Productive verbs facilitate action prediction in toddlers. *Infancy*, *19*(3), 301–325. https://doi.org/10.1111/infa.12047

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences*, *12*(10), 397–403. https://doi.org/10.1016/j.tics.2008.07.003

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences*, *95*(25), 15061–15065.

Hari, R., & Salmelin, R. (1997). Human cortical oscillations: A neuromagnetic view through the skull. *Trends Neurosci*., (20), 44–49.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307. https://doi.org/10.1016/S0896-6273(03)00838-9

James, K. H., & Maouene, J. (2009). Auditory verb perception recruits motor systems in the developing brain: An fMRI investigation. *Developmental Science*, *12*(6), F26–F34. https://doi.org/10.1111/j.1467-7687.2009.00919.x

Kilner, J. M. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences*, *15*(8), 352–357. https://doi.org/10.1016/j.tics.2011.06.005

Kilner, J. M., Bott, L., & Posada, A. (2005). Modulations in the degree of synchronization during ongoing oscillatory activity in the human brain. *European Journal of Neuroscience*, *21*(9), 2547–2554. https://doi.org/10.1111/j.1460-9568.2005.04069.x

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007a). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166. https://doi.org/10.1007/s10339-007-0170-2

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007b). The mirror-neuron system: A bayesian perspective. *Neuroreport*, *18*(6), 619–623. https://doi.org/http://doi.org/10.1097/WNR.0b013e3281139ed0

Lam, K. J. Y., Bastiaansen, M. C. M., Dijkstra, T., & Rueschemeyer, S. A. (2016). Making sense: Motor activation and action plausibility during sentence processing. *Language, Cognition and Neuroscience*, *0*(0), 1–11. https://doi.org/10.1080/23273798.2016.1164323

Liao, Y., Acar, Z. A., Makeig, S., & Deak, G. (2015). EEG imaging of toddlers during dyadic turn-taking: Mu-rhythm modulation while producing or observing social actions. *NeuroImage*, *112*, 52–60. https://doi.org/10.1016/j.neuroimage.2015.02.055

Liew, S.-L., Sheng, T., Margetis, J. L., & Aziz-Zadeh, L. (2013). Both novelty and expertise increase action observation network activity. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00541

Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science*, *24*(4), 279–284. https://doi.org/10.1177/0963721415570732

Marshall, P. J., Bar-Haim, Y., & Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *113*(8), 1199–1208. https://doi.org/10.1016/S1388-2457(02)00163-3

Marshall, P. J., & Meltzoff, A. N. (2014). Neural mirroring mechanisms and imitation in human infants. *Phil. Trans. R. Soc. B*, *369*(1644), 20130620. https://doi.org/https://doi.org/10.1098/rstb.2013.0620

Moreno, I., de Vega, M., & León, I. (2013). Understanding action language modulates oscillatory mu and beta rhythms in the same way as observing actions. *Brain and Cognition*, *82*(3), 236–242. https://doi.org/10.1016/j.bandc.2013.04.010

Nyström, P., Ljunghammar, T., Rosander, K., & von Hofsten, C. (2011). Using mu rhythm desynchronization to measure mirror neuron activity in infants: Measuring mirror

neuron activity in infants. *Developmental Science*, *14*(2), 327–335. https://doi.org/10.1111/j.1467-7687.2010.00979.x

Palmer, C., Zapparoli, L., & Kilner, J. M. (2016). A new framework to explain sensorimotor beta oscillations. *Trends in Cognitive Sciences*, *20*(5), 321–323. https://doi.org/10.1016/j.tics.2016.03.007

Pfurtscheller, G. (1992). Event-related synchronization (ERS): An electrophysiological correlate of cortical areas at rest. *Electroencephalography and Clinical Neurophysiology*, *83*(1), 62–69. https://doi.org/10.1016/0013-4694(92)90133-3

Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. *Vision Research*, *41*(10–11), 1257–1260. https://doi.org/10.1016/S0042-6989(00)00235-2

Potter, D. M., & Griffiths, D. J. (2006). Omnibus permutation tests of the overall null hypothesis in datasets with many covariates. *Journal of Biopharmaceutical Statistics*, *16*(3), 327–341. https://doi.org/10.1080/10543400600609585

R Core Team. (2017). *R: A language and environment for statistical computing*. *R Foundation for Statistical Computing*. Vienna, Austria. Retrieved from https://www.R-project.org/

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*(2), 339–349. https://doi.org/10.1111/j.1467-7687.2008.00786.x

Rueschemeyer, S. A., Brass, M., & Friederici, A. D. (2007). Comprehending prehending: Neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience*, *19*(5), 855–865. https://doi.org/10.1162/jocn.2007.19.5.855

Rueschemeyer, S. A., Lindemann, O., van Rooij, D., van Dam, W., & Bekkering, H. (2015). Effects of intentional motor actions on embodied language processing. *Experimental Psychology*. Retrieved from http://econtent.hogrefe.com/doi/full/10.1027/1618-3169/a000031

Southgate, V., Johnson, M. H., Osborne, T., & Csibra, G. (2009). Predictive motor activation during action observation in human infants. *Biology Letters*, rsbl20090474. https://doi.org/10.1098/rsbl.2009.0474

Springer, A., Huttenlocher, A., & Prinz, W. (2012). Language-induced modulation during the prediction of others' actions. *Psychological Research*, *76*(4), 456–466. https://doi.org/10.1007/s00426-012-0411-6

Springer, A., & Prinz, W. (2010). Action semantics modulate action prediction. *The Quarterly Journal of Experimental Psychology*, *63*(11), 2141–2158. https://doi.org/10.1080/17470211003721659

Szagun, G., Stumper, B., & Schramm, S. A. (2009). *Fragebogen zur frühkindlichen Sprachentwicklung (FRAKIS)*. Frankfurt am Main: Pearson Assessment & Information GmbH.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., … Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*(2), 273–281. https://doi.org/10.1162/0898929053124965

Tomasino, B., Weiss, P. H., & Fink, G. R. (2010). To move or not to move: Imperatives modulate action-related verb processing in the motor system. *Neuroscience*, *169*(1), 246–258. https://doi.org/10.1016/j.neuroscience.2010.04.039

van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy. *NeuroImage*, *43*(4), 808–814. https://doi.org/10.1016/j.neuroimage.2008.07.057

Yoo, K. H., Cannon, E. N., Thorpe, S. G., & Fox, N. A. (2016). Desynchronization in EEG during perception of means-end actions and relations with infants' grasping skill.

British Journal of Developmental Psychology, *34*(1), 24–37.

https://doi.org/10.1111/bjdp.12115